



①9 BUNDESREPUBLIK  
DEUTSCHLAND



DEUTSCHES  
PATENT- UND  
MARKENAMT

⑫ **Offenlegungsschrift**  
⑩ **DE 198 04 047 A 1**

⑤① Int. Cl.<sup>6</sup>:  
**G 10 L 5/00**

②① Aktenzeichen: 198 04 047.4  
②② Anmeldetag: 3. 2. 98  
④③ Offenlegungstag: 5. 8. 99

DE 198 04 047 A 1

⑦① Anmelder:

DeTeMobil Deutsche Telekom MobilNet GmbH,  
53227 Bonn, DE; Deutsche Telekom AG, 53113  
Bonn, DE

⑦② Erfinder:

Kauschke, Ulrich, Dr., 48165 Münster, DE; Rast,  
Herbert Roland, Dipl.-Ing., 53127 Bonn, DE; Runge,  
Fred, 15838 Wünsdorf, DE

⑤⑥ Entgegenhaltungen:

DE 40 24 890 A1  
DE 28 20 645 A1

**Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen**

Prüfungsantrag gem. § 44 PatG ist gestellt

⑤④ Verfahren und Einrichtung zur Erhöhung der Erkennungswahrscheinlichkeit von Spracherkennungssystemen

⑤⑦ Die Erfindung betrifft ein Verfahren zur Erhöhung der Erkennungswahrscheinlichkeit von Spracherkennungssystemen. Das Verfahren beruht darauf, daß nach Eingabe eines zu erkennenden Begriffs wahlweise ein Nachtrainieren des gleichlautenden bereits abgespeicherten Begriffs erfolgt. Damit wird erreicht, daß die sprecherabhängige Erkennungsrate auch in akustisch gestörten Umgebungen verbessert wird.

BEST AVAILABLE COPY

DE 198 04 047 A 1

## Beschreibung

Die Erfindung betrifft ein Verfahren und eine Einrichtung zur Erhöhung der Erkennungswahrscheinlichkeit von Spracherkennungssystemen gemäß dem Oberbegriff des Patentanspruchs 1.

Spracherkennungssysteme werden heute vor allem in Computern, Kommunikationssystemen und anderen technischen Geräten eingesetzt, wo es auf leichte Bedienbarkeit oder schnelle Dateneingabe ankommt. Die bisher bekannten Systeme sind jedoch nicht ausgereift und arbeiten fehlerhaft, insbesondere wenn sie in akustisch gestörten Umgebungen betrieben werden. Ein zu erkennendes Wort wird dann oft falsch oder gar nicht erkannt. Dies erfordert vom Benutzer ein mehrfaches Wiederholen des zu erkennenden Wortes, was bei einem häufigen Auftreten eines Erkennungsfehlers zu unzumutbaren Verzögerungen führt.

Der vorliegenden Erfindung liegt deshalb die Aufgabe zugrunde, ein Verfahren und eine Einrichtung zur Spracherkennung derart weiterzubilden, daß die sprecherabhängige Erkennungswahrscheinlichkeit, insbesondere in akustisch gestörten Umgebungen, erhöht wird.

Diese Aufgabe wird durch die kennzeichnenden Merkmale des Patentanspruchs 1 gelöst.

Vorteilhafte Abwandlungen und Weiterbildungen der Erfindung sind Gegenstand der abhängigen Patentansprüche.

Erfindungsgemäß wird für Spracherkennungssysteme ein Nachtrainieren der Sprachmuster neu eingegebener, möglicherweise falsch erkannter oder gar nicht erkannter Begriffe vorgeschlagen. Nachtrainieren bedeutet, daß ein falsch oder nicht erkannter Begriff nach wiederholter Eingabe nicht einfach überschrieben wird, sondern mit den zuvor eingegebenen Begriffen verglichen und korreliert bzw. durch ein neues Muster ergänzt wird, um so für die Erkennung des Begriffs unwichtige Muster oder Nebengeräusche zu reduzieren oder auszublenden. Die Erfindung ist insbesondere für den Einsatz in Spracherkennungssystemen gedacht, die in akustisch gestörten Umgebung arbeiten, z. B. in Mobilfunkendgeräten, Telefonen etc.

Bei erfolgreicher Erkennung eines bereits als Sprachmuster abgelegten Begriffs wird das abgelegte Sprachmuster mit dem neu aufgezeichneten nachtrainiert. Diese Nachtrainieren erfolgt bei erfolgreicher Erkennung eines jeden bereits aufgezeichneten Begriffs  $n$ -mal, wobei die Zahl der Durchläufe frei einstellbar und jederzeit veränderbar ist. Durch dieses Nachtrainieren wird eine stetige Verringerung des Einflusses von variablen Nebengeräuschen auf das eigentliche, konstante Sprachmuster eines Wortes erreicht.

Gemäß einer Weiterbildung der Erfindung erfolgt bei unsicherer Erkennung eines Wortes eine erneute Eingabeaufforderung, die zum Nachtrainieren genutzt wird. Auch hier ist die maximale Zahl der Durchläufe frei einstellbar und jederzeit veränderbar. Wenn der Vergleich eines eingegebenen Sprachmusters mit abgelegten Sprachmustern eine ähnliche und wenig differenzierte Wahrscheinlichkeit der Erkennung für mehrere abgelegte Begriffe ergibt, werden dem Systembenutzer diese Begriffe vorgespielt und er wird aufgefordert, den anfangs eingegebenen Begriff nochmals zu sprechen. Bei dann erfolgreicher Erkennung wird das Nachtrainieren abgebrochen. Aus Gründen der Übersichtlichkeit wird bevorzugt, bei unsicherer Erkennung die Anzahl der vom System ausgegebenen in Frage kommenden Begriffe auf eine vorgegebene Anzahl zu begrenzen und die Wiederholung des Vorgangs auf z. B. drei unmittelbare Wiederholungen zu begrenzen.

Eine andere Weiterbildung sieht vor, daß bei einem Neueintrag eines Wortes bzw. Begriffs im System, nach Eingabe des neuen Begriffs, ein Vergleich mit bereits gespeicherten

Begriffen erfolgt. Das Spracherkennungssystem wird also dazu genutzt, um einen neuen Begriff mit bereits abgelegten Begriffen zu vergleichen und festzustellen, ob sich das Sprachmuster des neuen Begriffs so weit von den Sprachmustern der bereits abgelegten Begriffe unterscheidet, daß keine Fehlerkennung bzw. unsichere Erkennung erwartet wird. Führt jedoch das neue Sprachmuster im Rahmen eines Korrelationsvergleichs nach einem festgelegten Kriterium einer unsicheren Erkennung zu einer starken Ähnlichkeit oder wahrscheinlichen Übereinstimmung mit bereits abgelegten Sprachmustern, so wird der Benutzer vom System optisch oder akustisch informiert und zu einer wiederholten Eingabe des neuen Begriffs aufgefordert, die dem Nachtrainieren dient. Die Anzahl der aufeinanderfolgenden Wiederholungen ist frei wählbar und jederzeit veränderbar.

Im folgenden wird die Erfindung anhand mehrerer Zeichnungsfiguren näher erläutert. Aus den Zeichnungsfiguren und der zugehörigen Beschreibung ergeben sich weitere Merkmale der Erfindung. Es zeigt:

Fig. 1 schematisches Ablaufdiagramm des erfindungsgemäßen Verfahrens am Beispiel eines in einem Mobilfunkgerät eingesetzten Spracherkennungssystems;

Fig. 2 schematisches Ablaufdiagramm des Verfahrens bei der Verarbeitung von Neueinträgen.

Anhand der Zeichnungsfiguren wird das erfindungsgemäße Verfahren in Verbindung mit einem Mobilfunkgerät mit Spracherkennungsfunktion näher beschrieben. Es wird von einem Mobilfunkgerät mit Spracherkennung und Sprachausgabe ausgegangen. Das Mobilfunkgerät besitzt ein Telefonbuch mit Namens-einträgen denen jeweils eine entsprechende Rufnummer zugeordnet ist. Durch Sprach-eingabe eines im Telefonbuch abgelegten Namens kann der Benutzer einen Wahlvorgang oder eine andere Aktion auslösen.

Gemäß Fig. 1 wird nach erfolgreicher Erkennung eines bereits als Sprachmuster im Telefonbuch abgelegten Namens das abgelegte Sprachmuster mit dem neu aufgezeichneten nachtrainiert. Bei unsicherer Erkennung eines Namens erfolgt eine erneute Eingabeaufforderung, die zum Nachtrainieren genutzt wird.

Die anfangs in Schritt 1 festgesetzte Variable  $k$  bzw.  $k_{\max}$  beschreibt die Anzahl bzw. maximale Anzahl der Trainingsdurchläufe.

Nach Aufforderung zur und Eingabe des Namens, gemäß Schritt 2, erfolgt in Schritt 3 zunächst ein Vergleich, ob die maximale Anzahl der Nachtrainingsdurchläufe erreicht ist. Ist dies der Fall, wird gemäß Schritt 4 die Operation abgebrochen. Ist die maximale Anzahl der Durchläufe nicht erreicht, erfolgt durch das Spracherkennungssystem in Schritt 5 ein Vergleich des eingegebenen Namens mit den bereits im Telefonbuch gespeicherten Namen. Wird der eingegebene Namen gemäß Schritt 6 sicher erkannt, so wird die vom Benutzer gewünschte Aktion ausgeführt, Schritt 7, z. B. ein Verbindungsaufbau zum gewünschten Gesprächspartner. Weiters erfolgt in Schritt 8 ein Nachtrainieren des entsprechenden Telefonbucheintrags mit dem zuletzt eingegebenen, als richtig erkannten Namen. Die Aktion ist beendet, Schritt 9.

Wird der eingegebene Namen jedoch nicht sicher erkannt, und ergibt der Vergleich eines eingegebenen Sprachmusters mit abgelegten Sprachmustern gemäß Schritt 6 eine ähnliche und wenig differenzierte Wahrscheinlichkeit der Erkennung für mehrere abgelegte Namen, dann wird der Systembenutzer informiert und ihm werden die dem eingegebenen Namen ähnlichsten Namen vorgespielt, Schritt 10. Der Benutzer wird in Schritt 11 aufgefordert, den anfangs eingegebenen Namen nochmals zu sprechen, Schritt 12. Der Zähler  $k$  wird um eins erhöht, gemäß Schritt 13, und die Routine

springt zurück zu Schritt 3 und wiederholt ab diesem Schritt.

Gemäß Fig. 2 erfolgt bei einem Neueintrag eines Namens in das Telefonbuch ein Vergleich mit den bereits gespeicherten Namen. Zunächst wird der Schleifenzähler k in Schritt 20 auf Null gesetzt.

Der Benutzer wird dann in Schritt 21 zur Eingabe des neuen Namens aufgefordert.

Dann wird in Schritt 22 überprüft, ob die Anzahl der Durchgänge den definierten Wert überschritten hat. Ist dies der Fall, wird der neue Namen gemäß Schritt 25 im Telefonbuch abgespeichert und das Verfahren beendet, Schritt 26.

Ist der Zähler kleiner als der vorgegebene Wert, so vergleicht das Spracherkennungssystem in Schritt 23 den neuen Namen mit bereits abgelegten Namen. Falls sich gemäß Schritt 24 das Sprachmuster des neuen Namens so weit von den Sprachmustern der bereits abgelegten Namen unterscheidet, daß keine Fehlerkennung bzw. unsichere Erkennung erwartet wird, so wird der neue Namen gemäß Schritt 25 im Telefonbuch abgespeichert und das Verfahren beendet, Schritt 26. Führt jedoch das neue Sprachmuster im Rahmen eines Korrelationsvergleichs nach dem festgelegten Kriterium einer unsicheren Erkennung zu einer starken Ähnlichkeit oder wahrscheinlichen Übereinstimmung mit bereits abgelegten Namen, so wird der Benutzer in Schritt 27 vom System optisch oder akustisch informiert und in Schritt 28 zu einer wiederholten Eingabe des neuen Namens aufgefordert. Der neue Namen kann nochmals eingegeben werden, Schritt 29. Die wiederholte Eingabe dient gemäß Schritt 30 dem Nachtrainieren. Die Anzahl der aufeinanderfolgenden Wiederholungen ist frei wählbar und jederzeit veränderbar. Nach jedem Durchlauf wird gemäß Schritt 31 der Zähler um eins erhöht, und die Routine springt zurück zu Schritt 21.

gespeicherten Begriffen erfolgt, um festzustellen, ob sich das Sprachmuster des neuen Begriffs so weit von den Sprachmustern der bereits abgelegten Begriffe unterscheidet, daß keine Fehlerkennung bzw. unsichere Erkennung erwartet wird.

7. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß wenn der neue Begriff im Rahmen eines Korrelationsvergleichs nach einem festgelegten Kriterium zu einer unsicheren Erkennung oder zu einer starken Ähnlichkeit oder wahrscheinlichen Übereinstimmung mit bereits abgelegten Begriffen führt, der Benutzer vom System optisch oder akustisch informiert und zu einer wiederholten Eingabe des neuen Begriffs aufgefordert wird, die dem Nachtrainieren dient.

8. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß die Anzahl der Durchläufe des Nachtrainings beliebig einstellbar sind.

Hierzu 2 Seite(n) Zeichnungen

#### Patentansprüche

1. Verfahren zur Erhöhung der Erkennungswahrscheinlichkeit von Spracherkennungssystemen, **dadurch gekennzeichnet**, daß nach Eingabe eines zu erkennenden Begriffs wahlweise ein Nachtrainieren des gleichlautenden bereits abgespeicherten Begriffs anhand des eingegebenen Begriffs erfolgt.
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß das Nachtrainieren auf einem Vergleich durch Korrelation des eingegebenen Begriffs mit dem gleichlautenden, bereits abgespeicherten Begriffs basiert, wobei die jeweils charakteristischen Sprachmuster ermittelt und abgespeichert werden.
3. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß bei unsicherer Erkennung des zu erkennenden Begriffs eine erneute Eingabeaufforderung erfolgt.
4. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß wenn ein Vergleich eines eingegebenen Begriffs mit den im Spracherkennungssystem abgelegten Begriffen eine ähnliche und wenig differenzierte Wahrscheinlichkeit der Erkennung für mehrere abgelegte Begriffe ergibt, dem Systembenutzer diese ähnlichen Begriffe angezeigt oder vorgespielt werden und er aufgefordert wird, den anfangs eingegebenen Begriff nochmals zu sprechen.
5. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß die Anzahl der angezeigten Begriffe beliebig vorwählbar ist.
6. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß bei einem Neueintrag eines Begriffs im Spracherkennungssystem, nach Eingabe des neuen Begriffs, ein Vergleich mit bereits

BEST AVAILABLE COPY

- Leerseite -

**THIS PAGE BLANK (USPTO)**

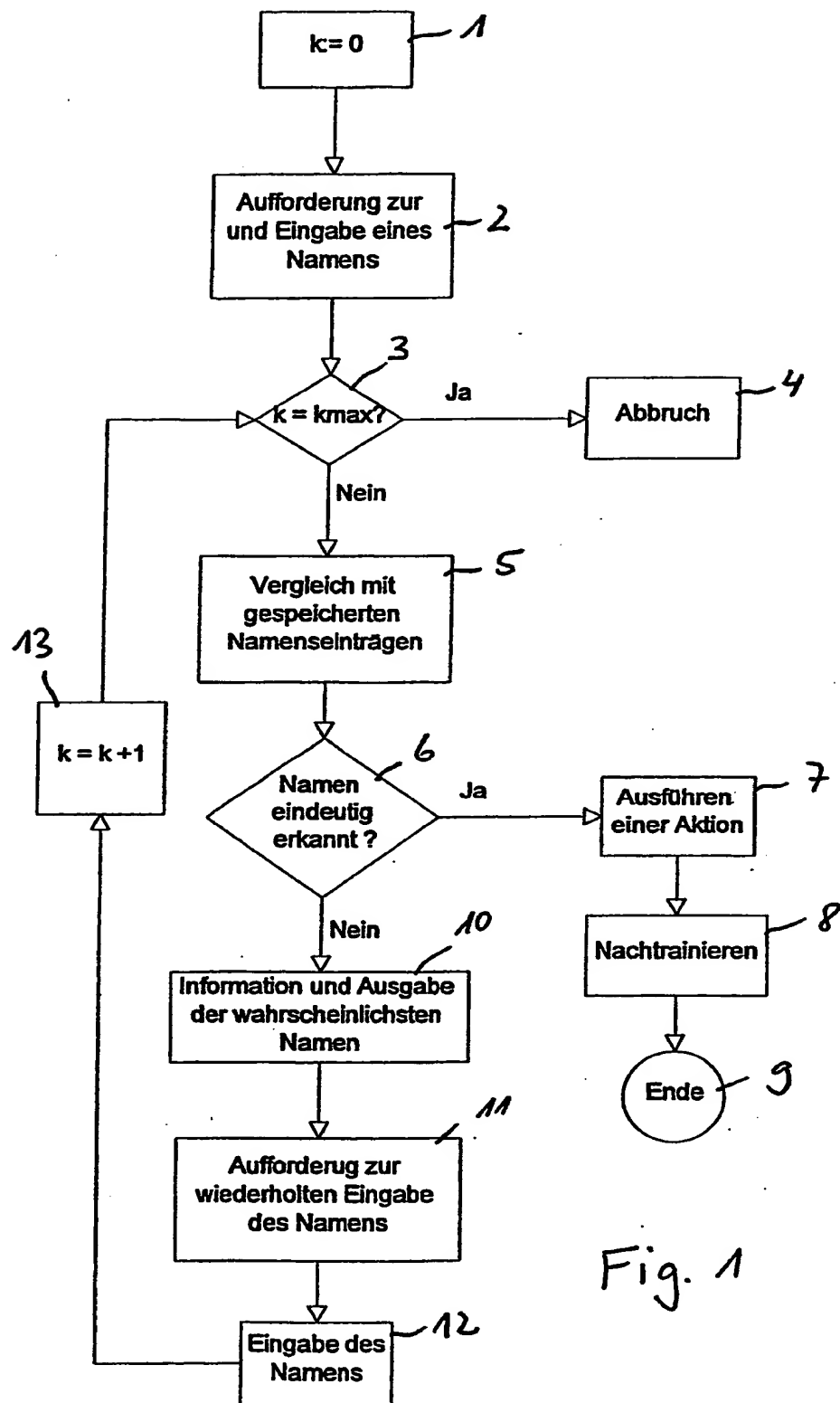


Fig. 1

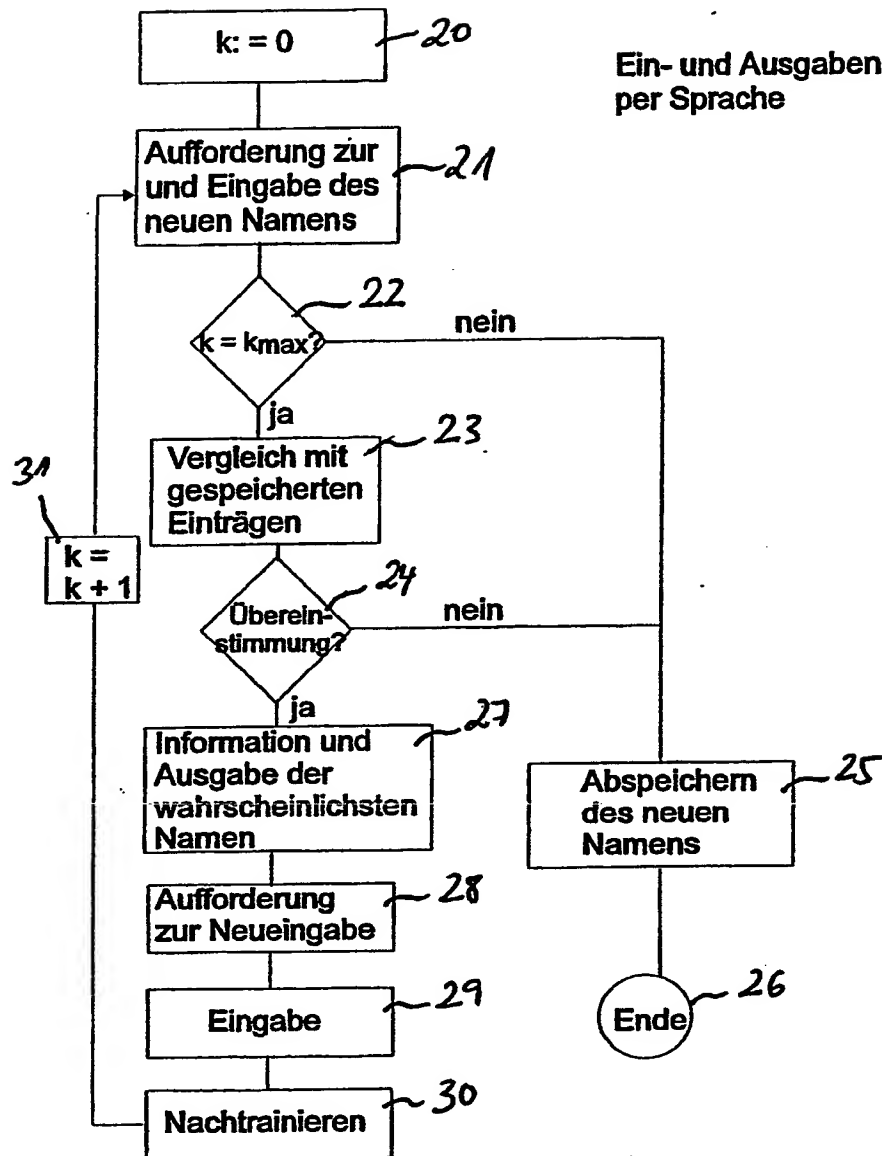


Fig. 2